Morton's Ranking of Races by Cranial Capacity

Author(s): Stephen Jay Gould

Source: *Science*, May 5, 1978, New Series, Vol. 200, No. 4341 (May 5, 1978), pp. 503-509

Published by: American Association for the Advancement of Science

Stable URL: https://www.jstor.org/stable/1746562

Morton's Ranking of Races by Cranial Capacity

Unconscious manipulation of data may be a scientific norm.

Stephen Jay Gould

Summary. Samuel George Morton, self-styled objective empiricist, amassed the world's largest pre-Darwinian collection of human skulls. He measured their capacity and produced the results anticipated in an age when few Caucasians doubted their innate superiority: whites above Indians, blacks at the bottom. Morton published all his raw data, and it is shown here that his summary tables are based on a patchwork of apparently unconscious finagling. When his data are properly reinterpreted, all races have approximately equal capacities. Unconscious or dimly perceived finagling is probably endemic in science, since scientists are human beings rooted in cultural contexts, not automatons directed toward external truth.

Nineteenth-century intellectuals discoursed endlessly on the subject of human racial differences; their works display an enormous excess of speculation based on a paucity of information. In pre-Darwinian America, polygenists argued for a separate (and unequal) creation of human races. Monogenists, placing their faith in scripture, traced all human diversity to an original Adam and Eve, and sought a scientific sanction for black inferiority in a greater degeneration from primeval perfection. Few Western scientists doubted the intrinsically higher status of their own white race, but opinion differed on the potential transience or innate permanence of black and Indian inferiority. Some approved slavery as the kindest status for lower races; others considered blacks inferior, but refused to justify slavery thereby. "Whatever be their degree of talents," wrote Thomas Jefferson (*1*), "it is no measure of their rights."

Morton the Objectivist

Samuel George Morton, a prominent Philadelphia physician, entered the melee, determined to replace idle speculation with hard fact. He set out to amass the world's largest collection of skulls, representing all racial groups (Fig. 1). He would, once and for all, obtain adequate samples to measure the physical differences among races. He began his collection in 1830 (*2*); it exceeded 1000 specimens when he died in 1851. More than 600 were sufficiently complete for an accurate account of cranial capacity—the most important physical measure of all, since Morton regarded it as a rough index of overall intelligence. (The general correlation of brain size and intelligence was not widely doubted in Morton's time.) Morton housed his collection—called "the American Golgotha" by his friends—at the Academy of Natural Sciences in Philadelphia, where he served as president from 1849 until his death.

Morton's collection was widely hailed as one of the wonders of the scientific world. Louis Agassiz wrote home to his mother about it (*3*): "Imagine a series of 600 skulls, mostly Indian, of all the tribes who now inhabit or formerly inhabited America. Nothing else like it exists elsewhere. This collection alone is worth a journey to America." Morton wrote at a time when American science was just beginning its transition from a stepchild of Europe to a vigorous enterprise worthy of attention and respect, even in the scientific centers of the Old World. America, Emerson wrote, had "listened too long to the courtly muses of Europe. . . . We will walk on our own feet; we will work with our own hands; we will speak our own minds" (*4, 5*).

Morton's work was hailed as a jewel of American science. Jules Marcou remarked that no zoologist except the great Cuvier had so influenced the thought of America's most illustrious scientific immigrant, Louis Agassiz (*5*, p. 102). On the occasion of Morton's death, the New York *Tribune* exclaimed that "probably no scientific man in America enjoyed a higher reputation among scholars throughout the world than Dr. Morton" (*5*, p. 144).

Morton did not achieve his reputation by astute interpretation or ingenuity of speculation—American science had been plagued by too high a ratio of theory to data. He won fame because he had finally presented a large body of objective fact. He had labored to collect and measure, where others had merely speculated. Oliver Wendell Holmes praised him for "the severe and cautious character" of his work, and for providing "permanent data for all future students of ethnology" (*6*). Europe's greatest scientific celebrity, Baron Alexander von Humboldt, wrote to Morton in 1844: "Your work is equally remarkable for the profundity of its anatomical views, the numerical detail of the relations of organic conformation, and the absence of those poetical reveries which are the myths of modern physiology" (*7*).

Morton's preference for data did not prevent him from holding opinions. He had a definite position and he defended it explicitly and often (*8–11*). As a prominent member of the polygenist school, he believed that the major human races had been created separately as true species. He argued that blacks and Caucasians were as distinct in ancient Egypt as they are today. Since humanity, following Moses, was not much more than 1000 years older than Egypt (*15*), races did not have enough time to differentiate from a common stock; they must have been created as we find them today. To the challenge that races interbreed freely and that sterility in crossing is the proper criterion of distinction, Morton replied by invoking both sides of the coin. Many true species hybridize and the traditional criterion must be revised (*9, 10*); offspring between some human races (Australoids and Caucasoids in particular) are both rare and deficient in fertility (*11*). But different need not mean unequal, and Morton needed a further criterion to defend the traditional ranking. Here he turned to his skulls, focusing almost exclusively on cranial capacity.

The author is a professor of geology and a member of the biology and history of science departments at Harvard University, Cambridge, Massachusetts 02138.

0036-8075/78/0505-0503$01.75/0

Morton published three major works on the cranial capacity of human races—the *Crania Americana* of 1839, a large, beautifully illustrated volume on Indian skulls (*12*); the *Crania Aegyptiaca* of 1844, his study on skulls from Egyptian tombs (*13*); and his summary of the entire Golgotha (623 skulls) in 1849 (*14*). Each of these works contained a summary table. These tables were frequently reprinted during the 19th century and became a linchpin in anthropometric arguments about human racial differences. Their supposedly objective hierarchies support, in detail, every Teutonic and Anglo-Saxon expectation for the ranking of races: whites on top, blacks on the bottom, and Indians in between; among Caucasians, Western Europeans on top, Jews in the middle, and "Hindoos" on the bottom.

The polygenist belief in a separate, created status for blacks and whites might have served as a primary defense for slavery in America; indeed, many polygenists (not including Morton) used their theory to support the South's "peculiar institution." But most apologists for slavery did not care to pay the price that polygeny demanded for its excellent argument—a denial of scriptural authority in the tale of Adam and Eve. After all, scripture can be bent to support any position, degeneration of blacks under the curse of Ham in this case. Darwin and Appomattox soon relegated the polygenic defense of slavery to oblivion, but Morton's hard data on cranial capacity survived as a cardinal input to any theory of racial ranking. In its obituary for Morton, the South's leading medical journal wrote: "We of the South should consider him as our benefactor, for aiding most materially in giving to the negro [sic] his true position as an inferior race" (*16*).

On Finagling Data

No scientific falsehood is more difficult to expunge than textbook dogma endlessly repeated in tabular epitome without the original data. Morton's tables enjoyed this brand of immortality and remained in the literature without serious challenge until the entire subject of racial ranking by cranial capacity fell into disrepute. But Morton, the self-proclaimed objectivist, did supply one rare and precious gift to later analysts: he published all his primary data with explicit statements on their genesis and manner of manipulation. We can learn exactly how he got from individual skulls to racial means.

I have reanalyzed Morton's data and I find that they are a patchwork of assumption and finagling, controlled, probably unconsciously, by his conventional a priori ranking (his folks on top, slaves on the bottom). The discrediting of some tables from the 1830's scarcely packs the punch of exposing Sir Cyril Burt's manipulation of data on IQ (*17*). I would regard this as a footnote to superannuated history if it did not raise so clearly a troubling issue that scientists usually sweep under the rug—and for good reason. I suppose that truly deliberate fraud to prevent the exposure of a suspected truth is rare in science. When we do uncover a case, we excommunicate its perpetrator, smugly declare that science purifies itself, and get back to work. Such cases rank high as gossip, but very low in telling us anything about the nature of normal, scientific activity. In fact, their hortatory value in the moralistic tradition permits us to avoid the issue; for we can pose our objective ideal against the transgression and pretend that the vast middle ground does not exist. However, I suspect that unconscious or dimly perceived finagling, doctoring, and massaging are rampant, endemic, and unavoidable in a profession that awards status and power for clean and unambiguous discovery. This is the middle ground of unappreciated bias and more conscious manipulation in the interest of a "truth" passionately held but inadequately supported.

Historians have occasionally studied this middle ground for insight into the genesis of creativity and the social constraints on scientific activity. We know, for example, that it has been occupied by many of our greatest heroes. Newton fudged outrageously to support at least three central statements that he could not prove (*18*). Any text in genetics will tell you that Mendel's F_2 ratios are too close to 3:1 to be believed. A kindly tradition, the Mendel's gardener hypothesis, attributes the finagling to a menial who knew what the boss wanted. But I can easily picture the good abbot himself, walking down a row of peas, a bit worried (in the absence of statistical knowledge) because his running tally stands at five tall plants too many, coming on a specimen, obviously tall but slightly below most of the others in stature, and saying to himself, "this one is not quite clear, so I'll skip it." The point is this: unconscious finagling is probably a norm. We need not protect the great by fobbing off responsibility for it on a laboratory assistant. We measure greatness not by "honesty," but by insight. After all, Newton and Mendel were right.

I do not want to sound flip. I do not condone or excuse finagling just because I regard much of it as intrinsic to scientific activity. I do share the scientist's faith that "correct" answers exist for most problems, and I believe that fudged data are paramount as impediments to solutions. I only raise what I regard as a pressing issue with two hopes for alleviation—first, that by acknowledging the existence of such a large middle ground, we may examine our own activity more closely; second, that we may cultivate, as Morton did, the habit of presenting

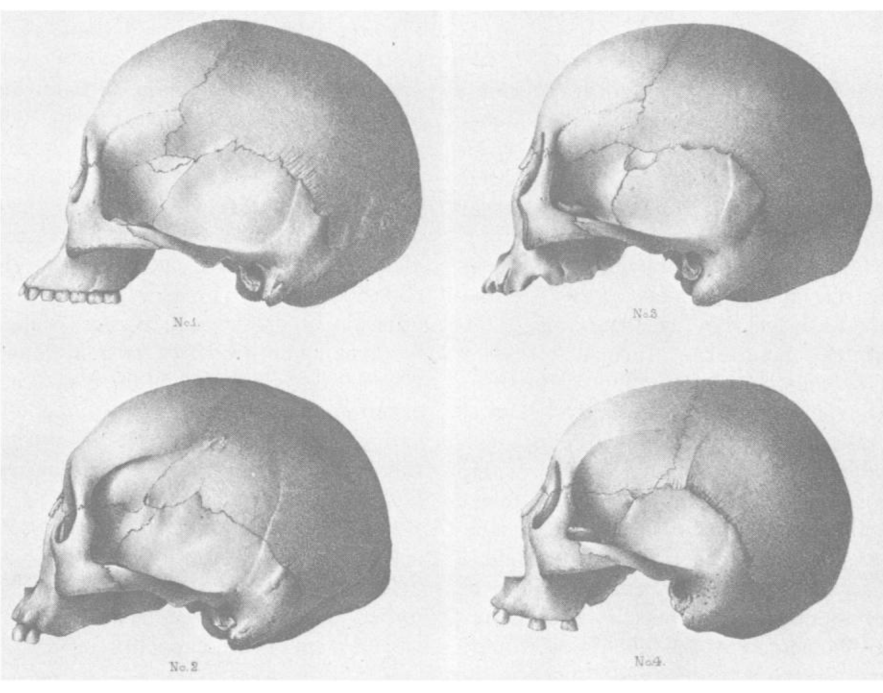

Fig. 1. Skulls of Eskimos. [Lithography by John Collins; printed in Morton's *Crania Americana* (*12*)]

candidly all our information and procedure, so that others can assess what we, in our blindness, cannot. But more general acknowledgment of the middle ground must come first. I suggest that some social scientists study the pervasive jokes, often self-directed, that scientists tell about finagling, and that others devise the most rigidly anonymous questionnaires.

In any case, since contemporary examples may be too threatening to inspire a general acknowledgment of the phenomenon, I present Morton on cranial capacity—an excellent example because the case is so distant and the controlling a priori so clear.

Crania Americana

"The benevolent mind," Morton concluded, "may regret the inaptitude of the Indian for civilization," but objective data had established it nonetheless, and sentimentality must yield to fact: "The structure of his mind appears to be different from that of the white man, nor can the two harmonize in their social relations except on the most limited scale" (*12*, p. 82). Morton had measured the capacity of 144 Indian skulls (*19*), calculated a mean of 82 cubic inches, 5 below the Caucasian average, and appended a table of phrenological measurements indicating a deficiency of "higher" mental power among Indians (*20*).

Morton began the *Crania Americana* with a dissertation on racial essences that discredits any claim to unbiased, dispassionate inquiry about the nature and meaning of human differences. The concept of "objective" knowledge is so culturally bound that Morton's supporters must have read these comments as evident truth, not Caucasian prejudice. For example, he wrote (*12*, p. 54): "Greenland esquimaux . . . are crafty, sensual, ungrateful, obstinate and unfeeling, and much of their affection for their children may be traced to purely selfish motives. They devour the most disgusting aliments uncooked and uncleaned. . . . Their mental faculties from infancy to old age, present a continued childhood. . . . In gluttony, selfishness and ingratitude, they are perhaps unequalled by any other nation of people." The "Hottentots," he wrote (*12*, p. 90), are "the nearest approximation to the lower animals. . . . Their complexion is a yellowish brown, compared by travellers to the peculiar hue of Europeans in the last stage of jaundice. . . . The women are represented as even more repulsive in appearance than the men." Yet, when Morton had to describe one Caucasian tribe as "a mere horde of rapacious banditti," he quickly added: "Their moral perceptions, under the influence of an equitable government, would no doubt assume a much more favorable aspect" (*12*, p. 19).

Table 1. Morton's summary table of cranial capacity by race (*14*, p. 260).

Race	N	Internal capacity (in^3) Mean	Largest	Smallest
Caucasian	52	87	109	75
Mongolian	10	83	93	69
Malay	18	81	89	64
American	147	82	100	60
Ethiopian	29	78	94	65

Turning to the central measures of cranial capacity, Morton's method is suspect from the start for two reasons. First, he did not distinguish male from female skulls. Since the mean sexual difference, due entirely to stature, is substantial (as we will learn from Morton's own data on Egyptian mummies), this failure is important—especially since many small subsamples contain skulls of one sex only. Second, he measured capacity by filling the skull with white mustard seed, sieved to reduce variation in grain size. But the seeds, by Morton's own later admission (*21*), were too light and still too variable in size to pack well, and the variation for remeasurements of the same skull ranged to 4 in^3. (Later, Morton switched to lead shot 1/8 in. in diameter "of the size called BB," and reduced the variation among measures of the same skull to less than 1 in^3.) Such an uncertainty will increase the variance, but it need not alter the mean for a series of skulls. It does, however, provide a wide berth for the influence of unconscious bias. Indeed, we know that Morton himself began to worry. He had hired assistants to measure the Indian crania (*21*), but, distressed by errors and inconsistencies, he later took to making all measurements himself (*14*) with lead shot.

Morton's Indian mean of 82 in^3 is a straight, ungrouped average of all skulls, representing Indian peoples from northern Canada to South America (Table 1). As a first observation of note, it is incorrect. He divides all Indians into two groups, the "Toltecans" from Mexico and South America, and the "Barbarous Tribes," largely from the United States and Canada. He gives a sample size of 147 (it should be 144 because three skulls were too incomplete for a measure of total capacity), 57 Toltecan and 87 Barbarous. However, he reports for the whole the Barbarous mean of 82.4 in^3 (rounded off in Table 1). Including the Toltecan mean of 76.8 in^3 and using his method of ungrouped averaging, the true grand mean is 80.2 in^3. (This elementary error permitted Morton to retain the conventional scale of being with whites on top, Indians in the middle, and blacks on the bottom.)

As a primary reason for rejecting Morton's ungrouped mean, I note widespread statistical inhomogeneity among his subsamples for various Indian peoples (Table 2). For example, $t = 8.47$ at 39 degrees of freedom, $P < .001$, for a comparison between Inca Peruvians ($N = 33$, $\bar{x} = 74.36$) and Seminole-Muskogees ($N = 8$, $\bar{x} = 88.28$). (Of course, we cannot fault Morton for ignoring a statistical procedure invented by Mr. Gosset of Guinness Breweries during this century. But I will show that Morton was well aware of problems posed by inhomogeneities among subsamples; in fact, they constitute the basis of his finagling.) It is intriguing that Morton often reported Caucasian means by sub-

Table 2. Means for Indian subsamples with more than four skulls.

People (Morton's characterization)	Mean measured with seed* (in^3)	N	Mean for same skulls measured with shot† (in^3)	N
Peruvians	74.4	33	76.6	33
Mexicans	80.2	13	82.5	9
Seminole-Muskogee	88.3	8	93.5	6
Western Lenapé	84.3	15	87.3	9
Northern Algonquin-Lenapé	88.8	4	91.3	4
Natick	79.7	9		<4
Osage	84.3	6	86.3	6
Iroquois	91.5	4		<4
Ohio Caves	84.9	9	87.6	5
Mounds	81.7	9	83.2	6
Mean	83.8		86.0	

*From *Crania Americana* (*12*). †From final catalog of 1849 (*34*).

samples, which permitted him to assert the superiority of Teutons and Anglo-Saxons. But he never broke down the Indian mean, even though he acknowledged a separate origin for several of the Indian peoples (*9*, p. 40). Thus, the fact that some Indian subsamples (Iroquois at 91.5 in^3, $N = 4$) exceeded the mean for Americans of Anglo-Saxon stock remained hidden in his raw data. (Morton never calculated the Indian subsample means at all; I have recovered them from his data.)

Morton's low mean of 80.2 in^3 reflects the accident of grossly unequal sample sizes. Inca Peruvians, with the smallest mean capacity ($\bar{x} = 74.36$) are most abundantly represented ($N = 33$, or 23 percent of the total sample). To weight Morton's subsamples equally, I computed the mean of means for all ten subsamples with more than four skulls (*22*). (Identification of subsamples comes from Morton's own tribal descriptions.) The mean capacity is 83.79 in^3 (Table 2).

This still leaves a large space between the Indian and Caucasian means. But we note that Morton's Caucasian sample of 52 purposely excludes 14 Hindu skulls for an interesting reason, openly stated (*12*, p. 261): "It is proper, however, to mention that but 3 Hindoos are admitted in the whole number, because the skulls of these people are probably smaller than those of any other existing nation. For example, 17 Hindoo heads give a mean of but 75 cubic inches; and the three received into the table are taken at that average." Thus, Morton was well aware that the sizes of subsamples can strongly and unfairly affect a mean—yet he included a large subsample of the smallest heads to pull down the Indian mean, and excuded just as many small Caucasian heads to raise the mean of his own group. Since he tells us what he did so explicitly, I must assume that he deemed his procedure proper. But by what rationale—unless it was the a priori assumption of a truly higher Caucasian mean? For then one might throw out the Hindu sample as truly anomalous, but keep an Inca subsample (with the same mean) as the lower end of normality for its disadvantaged larger group.

We, in any case, must follow our procedure of weighting all subsamples equally. The Caucasian sample represents four of the "families" that Morton included in the group. We cannot reconstruct the family means, since most skulls are labeled as "Europeans, nation not ascertained," but we can at least ensure that Hindu skulls constitute one-fourth of the total. If we restore the 14 Hindu heads that Morton excluded, we have 17 Hindu skulls in a total sample of 66, or 26 percent of the total. The Caucasian mean is now 84.45 in^3 (average of 52×87 reported by Morton and 14×75 for the added Hindus). Thus, from a great disparity between 80.2 in^3 for Indians and 87 in^3 for Caucasians, we recalculate a fairest estimate of 83.79 in^3 for Indians and 84.45 in^3 for Caucasians, or no difference worth mentioning. (Eskimos, despite Morton's low opinion of them, give a mean of 86.8 in^3, hidden by amalgamation with other subgroups in the Mongol grand mean.)

Table 3. Cranial capacity of Indian groups ordered by Morton's assessment of body stature. I have amalgamated some subsamples into Morton's larger tribal groups. Hence some groups that are not in Table 2 appear here. Morton did not include Columbia River Flatheads in his mean because the crania are distorted according to tribal customs for shaping the head. Morton states, however, that flattening distorts proportions but does not alter the mean capacity.

Stature and group	Cranial capacity (in^3)	N
Large		
Seminole-Muskogee	88.3	8
Chippeway and related groups	88.8	4
Dacota and Osage	84.4	7
Middle		
Mexicans	80.2	13
Menominee	80.5	8
Mounds	81.7	9
Small		
Columbia River Flatheads	78.8	10
Peruvians	74.4	33

We are still left with large differences among subgroups of both Indians and Caucasians (although a similar range of subgroup means for both). Why are Inca Peruvians low and Seminoles high, a fact that bothered Morton considerably when he considered the splendors of the Inca empire—although he consoled himself with the ease and rapidity of their defeat by the conquistadors. From allometric studies, we know that body stature is the primary determinant of differences in brain size among human groups, sexes, or races (*23*). Since Hindus are by far the smallest of Morton's Caucasian peoples, we may expect a similar correlation for Indians. Morton gives no hard data on stature, but his descriptions of some tribes do permit a rough division into small, medium, and large (I merely repeat Morton's assessment to show that he might have seen the correlation himself, had he been looking for it; I do not vouch for its accuracy.) Table 3 presents this division for all groups with more than four skulls. The correlation of brain and body size is affirmed without exception. We have no reason to attribute Morton's cranial differences among subsamples to anything other than variation in average body size.

Crania Aegyptiaca

Morton's study of mummified remains led him to the gratifying conclusion that the wonders of ancient Egypt had been designed by Caucasians. Blacks were present, as distinct from whites at the dawn of human history as they are today—a powerful argument for separate creation. "Negroes," Morton writes, "were numerous in Egypt, but their social position in ancient times was the same that it now is, that of servants and slaves" (*13*, p. 158).

Morton appended the following interesting footnote to his summarized table of cranial capacity (*13*, p. 113): "I have in my possession 79 crania of Negroes born in Africa. . . . Of the whole number, 58 are adult, or 16 years of age, and upwards, and give 85 cubic inches for the average size of the brain. The largest head measures 99 cubic inches; the smallest but 65. The latter, which is that of a middle-aged woman, is the smallest adult head that has hitherto come under my notice." I have two comments on this.

1) An addition of 29 skulls to the 1839 sample of 29 raised the mean by 6 in^3 to a value *above* the properly readjusted Caucasian mean of 84.45 in^3 and not far below Morton's own value of 87 in^3. Surely something funny is going on here. If the 1839 mean of 78 in^3 is correct (see Table 1), then the average capacity of the new skulls must be 92 in^3 to raise the grand mean to 85 in^3.

I suspect instead the change in method from mustard seeds to lead shot; the lighter mustard seeds did not pack well, leaving empty space in a "filled" cranium and giving a systematically lower capacity than that obtained with shot. Fortunately, we can test for differences because Morton personally remeasured all his skulls with shot and recorded the values in his final catalog (*24*). For 111 Indian skulls, 92 give higher values for shot than for seed. The average increase per skull (for all 111 skulls) is 2.2 in^3.

Unfortunately, Morton did not specify African and Caucasian skulls individually in his 1839 monograph; moreover, he borrowed some skulls from friends and included data from other sources in computing the black and white means. These skulls were never remeasured with shot. Still, we can make some inferences about systematic bias in the original measurements with seed (I will assume, as Morton contends, that measurements with shot were objective and invariably repeatable to within 1 in^3). Morton remeasured 18 of 29 African skulls from *Crania Americana*. With shot, they give a mean of 83.44 in^3, or an average rise of 5.4 in^3 from the 1839 mean of 78 in^3. Is this difference between African and Indian corrections (5.4 versus 2.2 in^3) an artifact of the incomplete African sample, or does it indicate a systematic undermeasurement of black skulls with the subjective method of mustard seeds? (I have presented other evidence of finagling to place blacks below Indians.) I strongly suspect a systematic bias for undermeasurement of black skulls. If the actual rise for all 29 skulls were, as for Indians, 2.2 in^3, then the 11 remaining African skulls would have a mean capacity with shot of 74.90 in^3, or 3.1 *below* the grand mean with seeds. Only 8 of 77 African skulls in Morton's final catalog have capacities below 74.9 in^3.

The data for Caucasian corrections are more ambiguous since only 19 of Morton's 49 European skulls were remeasured to appear in his final catalog. Removing the 3 Hindu skulls (since we cannot tell which ones he chose) from the 1839 sample of 52 Caucasians, the mean for the remaining 49 is 87.73 in^3. Nineteen non-Hindu Caucasian skulls remeasured with shot give 89.53 in^3, for an average correction of only 1.8 in^3 as a best estimate. The order of increasing correction for the switch from a subjective to an objective method matches the expected bias of desired underestimation: white, Indian, black.

2) Morton reported falsely that the smallest black skull was the smallest among all people that he had ever seen (*25*). But three Inca Peruvian skulls are recorded as 60, 62, and 64 in^3 in *Crania Americana*. Remeasured with lead shot, four skulls of this original series are smaller: 58, 62, 62, and 63 in^3 (*26*). Five additional Peruvian skulls measure less than 65 in^3 in Morton's final catalog (*27*). Again, I cannot get inside his ample head, but I suspect an a priori desire to keep blacks at the bottom as an impetus to amnesia.

The summary table of ancient skulls from the Egyptian tombs (Table 4) confirms every Western European's desire. Among Caucasians, Pelasgics (Hellenes, or ancient Greek forebears) exceed Jews and Egyptians. Negroids (mulattoes with more Negro than Caucasian features) are next, and pure blacks are last.

Morton's subdivision among Caucasians represents a false, typological breakdown of continuous variability (with ethnographically incorrect assignments as well). It should be ignored and the samples amalgamated to give a Caucasian mean of 82.15 in^3 [$N = 65$, standard distribution (s_x) = 7.76], well below the modern black mean. If we give Morton the benefit of the doubt anyway, and rank his three subsamples equally, we get a mean of 83.3 in^3 [(88 + 82 + 80)/3]. Still, this exceeds substantially the Negroid and Negro means.

Table 4. Cranial capacities for skulls from Egyptian tombs (*13*, p. 113).

People	Mean capacity (in^3)	N
Caucasian		
Pelasgic	88	21
Semitic	82	5
Egyptian	80	39
Negroid	79	6
Negro	73	1

But if we put Morton's subjective divisions aside, and separate Caucasians into male and female (sexual determinations could be made on many of these mummified remains), we obtain the following remarkable result. For 24 skulls, identified by Morton as male, $\bar{x} = 86.46$ ($s_x = 6.61$; range, 76 to 97 in^3). Twenty-two female skulls give only 77.23 in^3 ($s_x = 6.38$; range, 68 to 90 in^3), for a difference of more than 9 in^3. Turning to the six Negroid skulls, Morton identified two as female (71 and 77 in^3). He was unable to determine sex for the other four (77, 77, 87, and 88 in^3). In his final catalog of 1849, Morton guessed at the sex (and age, to within 5 years) for nearly all his crania. Here he specified the crania measuring 77, 87, and 88 in^3 as male and the other 77-in^3 skull as female—for a male mean of 84.0 and a female mean of 75.0 in^3, or 2.5 and 2.2 in^3 lower than Caucasian means by sex. But suppose that the two 77 in^3 skulls are female, and the 87 and 88 in^3 male (this hypothesis is just as likely since clean skulls cannot be identified unambiguously by sex, as Morton realized when he declined to specify in his original work). Then the male Negroid mean would be 87.5 in^3, slightly above the Caucasian male mean, while the female Negroid mean of 75.5 in^3 would be slightly below the Caucasian. The apparent difference of 4 in^3 between grand means for Negroids and Caucasians would only reflect the fact that about half the Caucasian sample is male, while only one-third of the Negroid sample may be male. (The apparent difference is magnified by Morton's incorrect rounding of the Negroid mean down to 79 rather than up to 80 in^3. As we will

Table 5. Morton's final summary of cranial capacity by race (*34*).

Races and Families	N	Cranial capacity (in^3)			
		Largest	Smallest	Mean	Mean
Modern Caucasian Group					
Teutonic Family					
Germans	18	114	70	90	92
English	5	105	91	96	
Anglo-Americans	7	97	82	90	
Pelasgic Family	10	94	75	84	
Celtic Family	6	97	78	87	
Indostanic Family	32	91	67	80	
Semitic Family	3	98	84	80	
Nilotic Family	17	96	66	80	
Ancient Caucasian Group					
Pelasgic Family	18	97	74	88	
Nilotic Family	55	96	68	80	
Mongolian Group					
Chinese Family	6	91	70	82	
Malay Group					
Malayan Family	20	97	68	86	85
Polynesian Family	3	84	82	83	
American Group					
Toltecan Family					
Peruvians	155	101	58	75	79
Mexicans	22	92	67	79	
Barbarous Tribes	161	104	70	84	
Negro Group					
Native African Family	62	99	65	83	83
American-born Negroes	12	89	73	82	
Hottentot Family	3	83	68	75	
Australians	8	83	63	75	

see again, all of Morton's minor numerical errors favor his a priori biases.) Thus, the entire case for a lower Negroid mean rests on the dubious identification of a single skull as male—and the difference of 2 to 2½ in^3 is insignificant in any case.

The large mean difference between sexes also affirms the primary correlation of brain size with stature. Most readers will have correctly divined by now that the single pure Negro skull is a female. In summary, Egyptian evidence does not support a difference in cranial capacity between blacks and Caucasians. Both groups are below the average of modern African blacks.

Final Summary of 1849

Morton's burgeoning collection included 623 skulls when he presented his final tabulation in 1849. Morton mused with pride on the largest set of such data ever compiled—"a novel and important contribution to Ethnological science," he proclaimed (*14*, p. 221).

Again, Morton presented the Caucasian distribution by "family," from Germanic to Hindu (Table 5). He cited the problems posed by unequal subsample sizes (conveniently ignored for Indians) in refusing to calculate a Caucasian grand mean: "No mean has been taken of the Caucasian race collectively, because of the very great preponderance of Hindu, Egyptian and Fellah skulls over those of Germanic, Pelasgic and Celtic families" (*14*, p. 223). First, his statement about a "very great preponderance" is false. Among modern Caucasians, $N = 46$ for Germanics, Pelasgics, and Celtics, while $N = 49$ for Caucasian families with smaller crania. If we amalgamate the modern crania with the ancient Egyptian ones, $N = 64$ for families with larger crania and 104 for smaller crania. If we weigh the six modern subsamples equally, the mean of subsample means gives a modern Caucasian average of 85.3 in^3. The ancient Caucasian grand mean for two families is 84.0 in^3.

Finally, all three means for Teutonic and Anglo-Saxon groups are incorrect or biased in Morton's favor. The German mean, reported at 90 in^3 in the summary, is 88.4 in^3 from individual skulls listed in the final catalog; the Anglo-American mean of 90 in^3 is really 89.14 in^3. The high English mean of 96 in^3 is correct, but the sample is entirely male (*28*).

Morton cites 82 in^3 for the Mongolian mean, based on a sample of six Chinese skulls. This low value reflects two examples of selective amnesia. First, Morton excludes the latest Chinese addition to his catalog (No. 1336 at 98 in^3), although he must have had the skull when he published his summary because many Peruvian skulls with higher numbers are included. The Chinese mean of all seven specimens is 84.14 in^3. Second, although Morton deplores the absence of Eskimos from his own collection (*24*, p. 12), he does not mention the three Eskimo skulls measured in *Crania Americana*. (These belonged to his friend George Combe and do not appear in Morton's catalog.) Morton never remeasured these skulls with shot, but their mustard seed average of 86.8 in^3 may have been several cubic inches too low. These two subsamples give a conservative Mongolian grand mean of 85 in^3.

Table 6. Corrected values for Morton's final tabulation.

People	Cranial capacity (in^3)
Native Americans	86
Mongolians	85
Modern Caucasians	85
Malays	85
Ancient Caucasians	84
Africans	83

Morton's Indian mean had plummeted to 79 in^3. But, again, this low value only records an increasing inequality of subsample size. Small-headed (and small-statured) Peruvians had formed 23 percent of the 1839 sample; they now made up nearly half the total sample (155 of 338 skulls). Using the previous criterion, I took all subsamples with more than four skulls (*29*), recomputed the means for skulls remeasured with shot (Table 2), and caluclated an Indian mean of 86.0 in^3 (the seed-to-shot correction of 2.2 in^3 matches exactly the recalibration based on all individuals).

We must drop Morton's Australoid family from the Negro mean because he wanted to assess the status of African blacks, and we no longer accept a close relationship between the two groups (dark skin is a convergent feature). We should also drop the Hottentot sample of three. They are very small in stature, and all three crania are female (*30*). Native and American-born blacks should be amalgamated to a single sample with a mean capacity between 82 and 83 in^3, but closer to 83.

Thus, we correct Morton's conventional "chain of being" to the following remarkable account (Table 6). There are *no* differences to speak of among Morton's races; all have means between 83 and 86 in^3. If Western Europeans sought their absolute superiority in the high mean of their subsample, I note that several unreported subsample means for Indian peoples are equally high and that all Teutonic and Anglo-Saxon means are biased or miscalculated. In any case, differences for subsamples within larger groups seem to rest on variations in body size alone.

Conclusion

Morton's finagling can be ordered in a few general categories:

1) Favorable inconsistencies and shifting criteria. As a favorite tool for adjustment, Morton chose to include or delete large subsamples in order to match grand means with a priori expectations. He included Inca Peruvians to reduce the Indian mean and excluded Hindus to raise the Caucasian mean. In 1849, he declined to calculate a Caucasian mean at all because he claimed (falsely) that subsamples with small crania dominated his total collection. He also chose to present or not to calculate subsample means in striking accord with desired results. He presented them for Caucasians to demonstrate the superiority of Teutons and Anglo-Saxons, but never calculated them for Indian subsamples with equally high values.

There are many other examples of shifting criteria among Morton's smaller works. In 1848, for example, he computed a Shoshonee Indian mean of 76 in^3 for a sample of three female skulls. He cared little for Shoshonees and used the low mean to discredit them further—even though he had praised Inca Peruvians with their even lower means (for a sample including males as well). Of the Shoshonees, he wrote (*31*):

> Heads of such small capacity and ill-balanced proportions, could only have belonged to savages; and it is interesting to observe such remarkable accordance between the cranial developments, and mental and moral faculties. Perhaps we could nowhere find humanity in a more debased form than among these very Shoshonees, for they possess the vices without the redeeming qualities of the surrounding Indian tribes; and even their cruelty is not combined with courage. . . . . A head that is defective in all its proportions must be almost inevitably associated with low and brutal propensities, and corresponding degradation of mind.

2) Procedural omissions that seem obvious to us. Morton was convinced before he began that differences in cranial capacity reflected innate mental ability. Once he finagled the "right" result, he regarded his work as complete. He did not consider alternative hypotheses, al-

though his own data stared him in the face. Thus, he arbitrarily divided a continuous spectrum of Caucasian variability into "higher" and "lower" subsamples, but never thought of computing means by sex, even though his Egyptian mummies provided this information objectively. And he never recognized the correlation between brain size and body stature, although his own data displayed it so clearly—variation among Indian peoples, Hottentots versus taller blacks, males versus females.

Average differences between the sexes are particularly striking. I already recorded 9 in^3 for Egyptian Caucasians. For Morton's largest sample of Inca Peruvians (*32*), males average 77.5 in^3 ($N = 53$), while females average 72.13 in^3 ($N = 61$). For Germans, males average 92.2 in^3 ($N = 9$), females 84.25 in^3 ($N = 8$). Moreover, Morton included several unisexual groups in his final tables, all to his advantage. His highest mean, for Englishmen, is based on an all-male group; his lowest, for Hottentots, on an all-female sample.

3) Slips. Two obvious errors seem hard to explain unless their conformity with expected results (both demoted blacks) provided so much satisfaction that Morton never thought of checking himself. Most curiously, after 200 pages of minute documentation, he reported his Indian mean incorrectly, as falling between blacks and whites, rather than at par with blacks. He stated repeatedly that two black crania had the smallest capacities among all his skulls, even though several Inca crania were smaller by his own tabulated measure.

4) Convenient omissions. Morton excluded a large Chinese skull and an Eskimo subsample in the 1849 tabulation of Mongolian capacity, thus reducing the grand mean below the Caucasian average.

5) Miscalculations. All miscalculations that I have detected are in Morton's favor. He rounded a Negroid Egyptian mean down to 79 in^3, rather than correctly up to 80 in^3. He cited means of 90 in^3 for Germans and Anglo-Saxons, but the correct values are 88 and 89 in^3.

Yet, through all this juggling, I find no indication of fraud or conscious manipulation. Morton made no attempt to cover his tracks, and I must assume that he remained unaware of their existence. He explained everything he did, and published all his raw data. All I discern is an a priori conviction of racial ranking so powerful that it directed his tabulations along preestablished lines. Yet Morton was widely hailed as the objectivist of his age, the man who would rescue American science from the mire of unsupported speculation.

I regard Morton's saga as an admittedly egregious example of a common problem in scientific work. Without a priori preferences, we would scarcely be human; and good science, as Darwin noted so often (*33*), collects data to test ideas. Science has long recognized the tyranny of prior preference, and has constructed safeguards in requirements of uniform procedure and replication of experiments. Gross flouting of procedure and conscious fraud may often be detected, but unconscious finagling by sincere seekers of objectivity may be refractory. The culprit in this tale is a naive belief that pure objectivity can be attained by human beings rooted in cultural traditions of shared belief—and a consequent failure of self-examination.

One may argue that lying with statistics is easier than fudging an experiment and that a direct intersection with contemporary politics makes for a more passionate a priori, but I think that most scientists pursue their private battles with as much ardor and as much at stake. I propose no cure for the problem of finagling; indeed, I write this article to argue that it is not a disease. The only palliations I know are vigilance and scrutiny.

References and Notes

1. Letter to H. Gregoire, 1809; quoted in T. F. Gossett, *Race: The History of an Idea in America* (Schocken, New York, 1965), p. 52.
2. Morton did no fieldwork himself. He collected his skulls by extensive and impassioned correspondence. The dangers to which many colleagues submitted themselves in collecting Indian skulls serve as a mark of Morton's reputation.
3. Cited in E. C. Agassiz, *Louis Agassiz: His Life and Correspondence* (Boston, 1886), vol. 2, pp. 409–429.
4. From an address of 1847; quoted in Stanton (*5*, p. 84).
5. W. Stanton, *The Leopard's Spots: Scientific Attitudes Towards Race in America* (Univ. of Chicago Press, Chicago, 1960), p. 84.
6. Letter to Morton, 27 November 1849; cited in Stanton (*5*, note 4, p. 96).
7. Letter of 17 January 1844; reprinted in C. D. Meigs, *A Memoir of Samuel George Morton, M.D.* (Collins, Philadelphia, 1851), p. 48.
8. S. G. Morton, *Trans. Am. Philos. Soc.* **9**, 93 (1844); *Proc. Acad. Nat. Sci. Phila.* **5**, 81 (1850).
9. ———, *Am. J. Sci.* **3**, 39 (1847).
10. ———, *ibid.*, p. 203.
11. ———, *Proc. Acad. Nat. Sci. Phila.* **5**, 173 (1851).
12. S. G. Morton, *Crania Americana or, A Comparative View of the Skulls of Various Aboriginal Nations of North and South America* (Remington, Philadelphia, 1839).
13. ———, *Crania Aegyptiaca: Observations on Egyptian Ethnography, Derived from Anatomy, History, and the Monuments*, originally published in (*8*).
14. ———, "Observations on the size of the brain in various races and families of man," *Proc. Acad. Nat. Sci. Phila.* **4**, 221 (1849).
15. Morton was not an antiscientific biblical idolater. He accepted geological evidence for the antiquity of the earth, and was, himself, a distinguished early American paleontologist (he described, for example, the fossils collected during the Lewis and Clark expedition). He accepted the Mosaic date for human creation because no fossil evidence then existed for earlier hominids.
16. R. W. Gibbs writing in the *Charleston Medical Journal*; quoted in Stanton (*5*, p. 144).
17. N. Wade, *Science* **194**, 916 (1976).
18. R. S. Westfall, *ibid.* **179**, 751 (1973).
19. Table 1 incorrectly lists a sample size of 147. Three of these skulls were incomplete and could not be measured for cranial capacity.
20. Morton gave cautious support to the doctrines of phrenology, but he did not practice the art himself, and made little use of it in his works. His assistant made the phrenological measurements on his Indian skulls, and Morton appended to his *Crania Americana* an essay on phrenology by his friend George Combe.
21. S. G. Morton, *Proc. Acad. Nat. Sci. Phila.* **1**, 7 (1841).
22. I chose four skulls as a minimum subsample size in order to give a probability near 5 percent that the group would represent a single sex only (unisexual groups are a primary source of misinformation in Morton's tables). If (as may well not be the case) Morton's skulls had an equal chance of being male or female, then only 1 in 16 samples of four skulls would form a unisexual group. Two groups that seem to have sample sizes larger than four in Morton's tables are excluded here: Morton cites capacities for five Cherokee skulls, but his final catalog (*24*) indicates that two of these skulls represent small children and should not have been included by Morton's own stated criteria. I also eliminate Columbia River "Flatheads." These skulls were artificially distorted, and Morton himself excluded them from his calculation of the Indian mean.
23. H. J. Jerison, *Evolution of the Brain and Intelligence* (Academic Press, New York, 1973); S. J. Gould, *Contrib. Primatol.* **5**, 244 (1975); P. V. Tobias, *Am. J. Phys. Anthropol.* **32**, 3 (1970); H. Pakkenberg and J. Voigt, *Acta Anat.* **56**, 297 (1964).
24. S. G. Morton, *Catalogue of Skulls of Man and the Inferior Animals* (Merrihew and Thompson, Philadelphia, 1849).
25. This was no momentary slip, for he repeated the claim several times in other publications; for example, *Proc. Acad. Nat. Sci. Phila.* **2**, 64 (1844).
26. Numbers 76, 497, 498, and 688 from Morton's final catalog (*24*). The skulls figured in *Crania Aegyptiaca* begin with number 795 in the final catalog.
27. Numbers 1418, 1419, 1424, 1461, and 1464.
28. I am reasonably certain that Morton and I are using the same skulls since the numbers tally perfectly. I exclude skulls listed as "idiot" because Morton reports that he did so as well. I include skulls marked "lunatic" because Morton does not mention their exclusion and because their inclusion makes the sample number tally perfectly with Morton's. In any case, the skulls of lunatics are not, on average, smaller than their group means. My Anglo-American sample consists of the following skulls from Morton's final catalog: 7, 14, 24, 45, 552, 899, and 1108. My German sample includes 37, 58, 434, 706, 1060, 1062 to 1066, 1187 to 1190, 1192, 1193, 1247, and 1249.
29. There are numerous inconsistencies in numbering between the *Crania Americana* and the final catalog of 1849. In addition, Morton used several skulls that he did not own in *Crania Americana*. These do not appear in the final catalog. Hence, sample sizes for skulls in the 1849 catalog that were also measured with seed in 1839 are reduced from Morton's 1839 tabulation (see Table 2).
30. Numbers 1107, 1244, and 1351 of the final catalog (*24*).
31. S. G. Morton, *Proc. Acad. Nat. Sci. Phila.* **4**, 75 (1848).
32. I emphasize that I do not assert the validity of Morton's sexual identifications. (I think that Morton guessed from the relative size and gracility of the skull in several cases.) I report these differences merely to point out that Morton could have discovered and evaluated this enormous influence of sex (as he determined it) if he had ever chosen to make calculations on this basis.
33. The most famous statement is in his letter to H. Fawcett, 18 September 1861: "How odd it is that anyone should not see that all observation must be for or against some view if it is to be of any service."